STATE OF CALIFORNIA PETE WILSON, *Governor*

DEPARTMENT OF INDUSTRIAL RELATIONS
DIVISION OF WORKERS' COMPENSATION
45 Fremont Street, 31st Floor
San Francisco, CA 94105

ADDRESS REPLY TO:
P.O. Box 420603
San Francisco, CA 94142

May 8, 1998

To: Interested Parties

From: Casey L. Young
Administrative Director

Re: What Do Injured Workers Think of Their Medical Care?

I am very pleased to release this report on the development of a survey to assess " What Do Injured Workers Think of Their Medical Care?" This survey was developed and tested by the University of California, Berkeley, Survey Research Center and the Division of Workers' Compensation. I am hopeful that the survey will provide a useful tool for expanding our understanding of the injured workers' prospective on medical care and outcomes after injury. Such information can be of tremendous value to claims administrators, employers, and health care providers- both in making decisions about health care purchasing, and in improving the quality of services for injured workers.

Many organizations have already expressed interest in using the DWC survey. We are also exploring the possibility of creating a mechanism for pooling information from various survey users to create benchmark data for patient satisfaction in California's workers' compensation system. DWC will host a meeting to discuss the use of the survey later this spring. If you are interested in participating, please call Ms. Kathy Dervin, at (415) 975-0734.

What Do Injured Workers Think of Their Medical Care?

A report on the development of a survey

May 1, 1998

Division of Workers' Compensation
Department of Industrial Relations

This report was prepared by James Wiley, Ph.D., Karen Garrett, and Yu-Teh Chang of the Survey Research Center, University of California, Berkeley, and Linda Rudolph, M.D., Kathy Dervin, M.P.H., and Neil Maizlish, Ph.D., of the Division of Workers' Compensation, California Department of Industrial Relations.

The authors gratefully acknowledge the many individuals who helped in the development of the questionnaire and in the administration of the pilot survey:

Nora Blay (Pacific Business Group on Health), Thomas Ash (Mullikin Medical Group), Julia Faucett, R.N, Ph.D. (UCSF School of Nursing), Jo Mason (Blue Cross of California), Rosie Mazzamuto and Tom Clark (Health Plan of the Redwoods), Mary Jane De Mille, R.N. (FHP-HCO-now PacifiCare), Lawrence Miller, M.D. (CCN), Alex Tspisis (International Association of Machinists and Aerospace Workers-Local 1788), Chuck Johnson (Pacific Bell), Vera Kawamura (Ca. Dept. of Corporations), Joan Lichterman (East Bay RSI-Support Group), Dorsey Hamilton (Comp ALERT), Bonnie Conrad, R.N. and Connie Chiulli, R.N. (Kasier Permanente), Julianne Sum,J.D. (UC Berkeley-LOHP), Debbie MacLyman C.O.H.N.(Health Management Centers West), Karen Berger, D.C. (Calif Chiropractors Assoc.), David Disher, M.D. (WOEMA). Mike Herberger (Safeway), Bernie Scalisi (Ross Stores, Inc.), Stephen Schuman, M.D. (Occupational Health Associates), Mouktar Moussad and Gideon Letz, M.D. (State Compensation Insurance Fund), David Michaels, Ph.D. (City University of New York), Kelly Kyes, Ph.D. (University of Washington, School of Public Health), Jim Ellenberger (AFL-CIO, Health and Safety Department), Mark Miller (CWCI), Glenn Pransky, M.D. (University of Massachusetts, Occupational Health Program), Tim Schmidle, Ph.D.(Cornell University- Institute for Labor and Industrial Relations**,** DWC staff (Glenn Shor, Ph.D., Vince Roux, M.B.A., Cynthia Robinson, M.P.P.) and UCB/SRC staff (Toni Genalo, Frank Newhauser, Alan Taniguchi).

Executive Summary

With the recent rapid changes in the health care delivery system, health care purchasers , consumers, and providers have become increasingly interested in assessing the quality of care. Several components of the quality of health care can best be assessed by patient reports on their experience; these include access to care, interpersonal aspects of medical care, and some aspects of outcomes of care. In the workers' compensation system, satisfaction with care may influence length of employer control over medical care, litigation rates, and return to work outcomes; workers' perceptions of outcomes after injury may also impact work productivity and morale. Information about worker's compensation patients' satisfaction with medical care and perceptions of outcome has, to date, been extremely limited.

The 1993 California workers' compensation reform legislation introduced Workers' Compensation Health Care Organizations (HCOs), intended to expand the use of managed care in California's workers' compensation system. The Division of Workers' Compensation (DWC) is mandated to certify HCOs, and to determine plan effectiveness. DWC contracted with the University of California, Berkeley, Survey Research Center (SRC) to develop a standardized self-administered questionnaire that could be used to collect data on patient satisfaction and outcomes in injured workers receiving care in HCOs and other health care delivery systems, and to conduct a pilot test of the survey instrument and mail-out procedure.

SRC and DWC reviewed available patient questionnaires used in other settings, and developed questions specifically oriented to occupational medicine and return-to-work outcomes. An ad hoc advisory committee reviewed a draft of the survey, and provided comments for questionnaire revision. The revised questionnaire was evaluated in a focus group with injured workers, which resulted in further cuts in length and revisions of specific questions.

The draft questionnaire was mailed to a sample of 800 workers, randomly selected from the claims files of six cooperating organizations; nearly 30% of workers returned the survey. SRC then conducted an intensive telephone follow-up to those workers who did not return the mail survey, obtaining a response rate of over 60%.

An analysis of characteristics and responses of both mail and phone respondents was conducted. Females and older workers were over-represented among mail respondents.

Phone respondents - who were interviewed at a later point in time relative to injury date - reported better outcomes, suggesting a “recovery effect”. There were no significant differences between mail and phone respondents with respect to satisfaction with medical care. SRC also assessed item non-response, validity of response, response variability, and item consistency.

Although the primary purpose of the pilot survey was to validate and finalize the questionnaire, several findings are of interest. Over one-quarter of respondents indicated that they were dissatisfied overall with the medical care received after a work injury. A quarter also reported that the physician did not understand the effect of the injury on their ability to perform their job. Only a minority of workers felt they had fully recovered by 6 months after the injury, and a large proportion reported continued pain from the injury.

These preliminary findings indicate the need for further investigation into patient satisfaction with workers’ compensation medical care, and patient perceptions of outcomes following work injury. Based on the SRC analysis, DWC has revised the survey instrument, which is now available for public use. Recommendations for administration of the worker survey, and plans for further implementation by DWC, are addressed.

I. INTRODUCTION: Goals of the Injured Worker Patient Satisfaction Project

In 1993, the California legislature enacted workers' compensation reform legislation which mandated the Division of Workers' Compensation (DWC) to certify managed care workers' compensation Health Care Organizations (HCOs)[1]. Workers enrolled in an HCO are required to obtain all medical care for a work injury from the HCO, for up to one year after the injury. HCOs are required to provide the information necessary for DWC to evaluate their effectiveness in providing care to injured workers; DWC's HCO Certification Standards require that HCOs "provide for periodic evaluation of the HCO by enrollees.... in the form and manner prescribed " by DWC.[2] Additionally, the reform legislation mandated DWC to establish a workers' compensation information system which would provide information about the adequacy of workers' compensation benefits received by injured workers. These mandates provide the context in which DWC began to address the issue of patient satisfaction with medical care in the workers' compensation system.

Patient satisfaction and patient perceptions of outcomes have become valuable and important components in the assessment of the quality of health care.[3] Patients are uniquely able to provide information about their ease or difficulty of obtaining care, the interpersonal dimensions of the patient-physician relationship, the patient's view of the technical quality of care provided, and the patient's functional status and perceived well-being.

State and federal government agencies (e.g. HCFA for Medicare, MediCal) and private and public purchasing groups (e.g. PERS, Pacific Business Group on Health) now require the collection of patient satisfaction data. However, there has often been little comparable information about patient satisfaction from plan to plan [4], and plans, purchasers, and individual consumers have all applied information from various surveys in different ways.[5] There has thus been a move toward standardization of patient or consumer surveys. For example, the National Council for Quality Assurance, HEDIS 3.0, includes a Member Satisfaction Survey which is widely used by health plans and some large purchasing coalitions[6] ; the Agency for Health Care Policy and Research developed the Consumer Assessment of Health Plans (CAHPS) which is intended to become a standardized patient assessment tool[3].

Early patient satisfaction surveys focused primarily on patient views of convenience of care, adequacy of facilities, interpersonal aspects of care, and overall satisfaction with care. Concurrently, other surveys were developed to assess population health status and functional capacity[7]. More recently, there have been efforts to incorporate both patient

satisfaction and patient perceived outcomes within a single survey; both HEDIS 3.0 and CAHPS do so.

Most patient satisfaction surveys neither collect nor report on the health care experiences of sub-groups (such as injured workers) within a population. Although several studies have looked at outcomes in injured workers with particular diagnoses and treatments[8-10], there had been only one prior assessment of patient satisfaction with workers' compensation medical care at time this project was initiated, and another state-based survey which was developed concurrently. Both were performed for the purpose of evaluating a specific, limited term managed care pilot program, rather than to provide on-going assessment of injured workers' perceptions of care [11, 12, 13].

The purpose of this project was to develop and test survey procedures for routinely obtaining standardized information about the experiences of injured workers in getting medical care for their injuries, and about injured workers' perceptions of well-being after a work injury. A full descriptive analysis of patient satisfaction and outcomes was beyond the scope of this project.

DWC contracted with the Survey Research Center (SRC) of the University of California, Berkeley, to:
a. develop a self-administered questionnaire that could be used to collect data on patient satisfaction and patient outcomes;
b. test a mail-out procedure for questionnaire administration that could be used by HCOs to fulfill DWC's requirements, and by other organizations to assess their patient or client population's satisfaction with medical care;
c. make recommendations for the analysis and reporting of survey data collected according to the mail-out procedure.

The challenges that routinely confront surveys of all types - issues of sample bias and response rates, problems of data accuracy and reliability, and limitations on the coverage of important issues related to the subject matter of the survey - are also present in research on patient satisfaction. Surveying patient evaluations and outcomes for injured workers presents an even more difficult challenge than surveys of patient satisfaction in group health contexts. This report describes some of the problems that were confronted in attempting to design and execute a survey of patient satisfaction among injured workers, and reports preliminary findings from that survey. Part II describes the methodology of the survey, including questionnaire design and coverage, sampling issues, and field procedures. Part III gives an account of the findings, with special attention to completion rates and response bias, questions of data quality, and analysis of subjective ratings of satisfaction with medical care. Part IV concludes by summarizing the implications of our experience for design and conduct of survey studies of injured workers' perceptions of their medical care and outcomes after work injury.

II. METHODS

1. Instrument Development

The goal of the questionnaire development phase of the project was to construct and test a self-administered questionnaire that could be used for mail-back surveys of injured workers receiving medical care through the workers' compensation system. Unlike patient questionnaires used in other settings, this questionnaire included a) broad coverage of patient outcomes (e.g. disability and pain), b) questions specifically aimed at the quality of occupational medicine, and c) questions pertaining to return-to-work and work modifications. The construction of the questionnaire used in the pilot survey proceeded in three stages:

Preliminary draft construction: DWC and SRC staff reviewed and discussed numerous questionnaires used to assess patient satisfaction in the context of group health, health status and disability questionnaires, and the few available questionnaires (including drafts) which had been constructed to look at injured worker patient satisfaction or outcomes. (See Appendix A for a list of all questionnaires reviewed). Additionally, we reviewed transcripts of prior focus groups at which injured workers discussed their medical care.[1]

Several issues bearing on questionnaire development surfaced during this process. First was the critical challenge of focusing the respondent on the medical care and outcomes which the survey is intended to evaluate. Most patient satisfaction questionnaires ask for relatively generic feedback about medical care. Questionnaires may distinguish between care received at the last visit and care in the prior year, between hospital and outpatient care, or between the respondents' views of medical care in general versus their own medical care. But few questionnaires distinguish between a woman's experience with her internist, her gynecologist, and her children's' pediatrician, although each of those contributes to her overall experience as reflected in questionnaire responses. Similarly, most population-based health status surveys ask about overall health status and functioning, rather than the impact of a particular health event on health and function.

Because the distinct purpose of this questionnaire is to assess patient satisfaction with workers' compensation, and patient perceptions of outcomes after work injury, this issue required considerable attention. The earlier focus groups suggested that workers need prompting to distinguish between care received in the workers' compensation system, and other medical care. Thus, we constructed several questions intended to initially focus the injured worker on workers' compensation medical care received for a specific work

[1] Focus groups were conducted by the UC Berkeley Labor Occupational Health Program, under contract with the California Industrial Medical Council, as part of a project to develop materials about medical evaluations for injured workers.

injury; we also modified questions drawn from other patient satisfaction surveys to repeatedly remind the respondent to focus on the care for that injury. Similarly, we modified questions drawn from health status surveys to elicit the worker's perception of how the specific work injury effected health and functional status.

Patient surveys provide unique information on what actually happens in the physician-patient interaction - information that cannot be derived from utilization or other outcomes data. However, none of the instruments reviewed included any questions specifically pertinent to occupational medicine. We therefore constructed several items to assess what experts consider to be several key aspects of the treatment of occupational injury and illness[14]. These include questions about the physician's effort to elicit information about the worker's job, whether the patient felt the physician understood the impact of the injury on ability to perform the job, and the extent of physician-patient communication about return-to-work.

Advisory Committee review: The preliminary draft instrument was distributed to a large number of individuals for comment. An ad-hoc advisory committee meeting was held at UC Berkeley in October, 1995; attendees included representatives of managed care organizations, labor unions, self-insured employers, workers' compensation carriers, and other potential users (see page 2 for a complete list of attendees and reviewers).

At the meeting, many reviewers commented about questionnaire wording, expressing concern that particular items would prompt either overly favorably or unfairly unfavorable responses. We revised several questions and attempted to ensure overall neutrality through wording changes and variable sequencing of responses (i.e. favorable response is not always first).

Reviewers raised concerns about numerous potential confounding variables that were not addressed in the draft instrument, and which could theoretically effect either satisfaction or outcomes. These included litigation, union membership, pre-existing illness and co-morbidities, and job satisfaction. There was a recommendation to include questions about waiting times in medical offices. We attempted to include these variables, while also addressing the committee's unanimous opinion that the survey was excessively long.

The advisory committee also expressed concern that the instrument be applicable more broadly than just for the evaluation of certified HCOs; they also pointed out that evaluation of HCOs can only be done in the context of comparable information about injured workers' experiences with other delivery systems. Thus, the final instrument is designed to be useful in assessing medical care for injured workers in all delivery settings, and for the benchmarking of patient satisfaction and outcome.

Focus group evaluation: A focus group of injured workers was held in February, 1996, in Oakland. The participants were East Bay residents injured in Fall, 1995, identified from

the claims records of a large workers' compensation insurance carrier. Nine of eleven invitees attended the session, which was conducted by DWC and SRC staff; each participant received an incentive payment of $50 to attend the 3-hour session.

During the focus group, participants were first asked to complete a 44 item questionnaire. The participants took between 25 and 60 minutes to complete the survey. At least one worker had difficulty reading the survey questions. Several others asked for assistance. SRC researchers then led a discussion of the survey, going through question by questions, seeking to determine areas of misunderstanding, lack of clarity, or other difficulties the workers had in completing the survey.

The discussion identified several potential difficulties:
a) Respondents were generally unclear about what was supposed to happen in the course of their workers' compensation claim. They expressed confusions about "who was who", and could not clearly identify the various individuals (e.g. various medical providers, claims adjusters, health plans, and others) with whom they had contact about their medical care or other benefits.
b) Respondents were asked to rate the care they received from "doctors and other health care providers". Several workers said they considered their physical therapist the main "doctor"treating them because the physical therapist explained their condition and spent more time with them than other practitioners. Others had changed doctors several times and were unclear which doctor they should be rating. The concept of the "primary treating physician"was confusing.
c) A sequence of return to work question s proved difficult for these workers. They asked "which return to work?" in cases with multiple periods of temporary disability, and "what do you mean by return to work?" In responding to a question about stage of recovery, a respondent stated that the doctor said she was well enough to go back to work, but that she didn't agree with this assessment.
d) Although at the time of the focus group there was no enrollment in workers' compensation Health Care Organizations, several workers believed they were enrolled in an HCO.

Although a number of these workers had been receiving benefits in the system for over a year, the general lack of knowledge these workers displayed about the workers' compensation system was striking, but consistent with prior studies.[15]

As a result of the focus group, questions about case managers and HCOs were eliminated, skip patterns were simplified, terms such as "return to work"and "health care provider"were clarified, some questions were changed to focus on "the person who provided most of your care", and the survey instrument was shortened.

The self-administered questionnaire used for the pilot mail-out survey is found in Appendix 3. The questionnaire represents a balance of newly constructed questions

(mostly pertaining to occupational medicine and return to work) and questions taken from other instruments that were modified for the purposes of our study. It is designed to be understood by English-speaking patients with at least an 8th grade reading level, and to be completed in less than 30 minutes. The questionnaire consists almost entirely of items with pre-coded responses or fill-ins. The majority of items fall into five substantive categories (see Table 1): a) patient ratings of and reports about workers' compensation medical care; b) post-injury health and functional status; c) return to work; d) utilization of medical services; and e) demographic and occupational characteristics of injured workers. Though initially intended for use by organizations participating in the HCO program, the questionnaire is designed so that it can be employed generally in survey-based studies of medical care for injured workers.

Table 1: **Summary of the content of the self-administered questionnaire used for the pilot mail-back survey**

Topic	Question Numbers
Injury and Post-Injury Health	
self-rated health	1
injury date and type	2,3AB
post-injury health outcomes	4,5,6,7ABC
emotional problems	10AB
pain	8 ABC
Patient Ratings of Care	
help managing pain	9
access and waiting time	13,14ABC
overall satisfaction	16
patient involvement	17A,18A
communication and thoroughness	20B(1)-(5)
occupational medicine	17BCD,19BCD,20B(6)
Utilization of Services for Injury	15AB,20A
Return to Work and Work Limits	26ABCD,27ABC,28,29,30ABC
Occupational and Demographic Characteristics of Injured Workers	23,33,34,35,36,37,38,39,40
Other	
other insurance coverage	11
out-of-pocket expense	21
use of vacation/sick leave	22
pre-injury job satisfaction	24,25
union member	31
attorney	32AB

2. Sampling

The sample used for the mail-back survey consists of a subset of injured workers selected from the claims files of cooperating organizations. A total of 2,651 eligibles were made available to SRC by six organizations: two self-insured employers (Ross Stores and Safeway), a large workers' compensation insurance carrier (State Compensation Insurance Fund), and three medical organizations (Kaiser-Permanente, Occupational Health Associates, and Mullikin Medical Group). These organizations provided a sampling frame with diversity in occupation, geographic location, source of workers' compensation medical care, and claims administration.

Claimants were eligible for selection into the survey if they were a) injured between September 1 and December 15, 1995, and b) had three or more days of lost time or received payment for temporary disability. The following information was requested for each case: name, address, home phone number, date of injury, type of injury, date of birth, gender, and date of release to return to work or end of temporary disability or actual return to work. Some difficulties were encountered in obtaining this information for all cases - particularly dates related to return to work. Type of injury was not provided in a uniform manner; some organizations provided ICD diagnoses, while others provided body part and/or injury category, using variable coding schemes.

Samples from each organization were selected at random. The sampling fractions differed for each organization, because of the wide variation in the number of cases contributed by the different organizations. For organizations submitting only a small number of claims, all eligible claimants were included in the sample; for large organizations which provided a large number of eligibles, the sampling fraction was lowered so that cases contributed by these organizations would not excessively dominate the sample.

The final sample size was set at 802 cases, in anticipation of obtaining between 200 and 300 completed questionnaires in the mail survey. The distribution of the sample with respect to type of contributing organization, nature of injury, and age and gender is shown in Table 2.

Table 2: **Distribution of sample of eligible survey participants by selected characteristics**

Characteristic	Distribution by Characteristic N=802 %
Type of Organization	
Self-insured	7.7
State fund	32.4
Medical organization	59.9
	100.0
Nature of Injury	
Back sprain/strain	34.2
Other sprain/strain	7.0
Other injury	58.9
	99.9
Age	
less than 30	35.7
31-50 years old	52.1
51 and over	12.2
	100.0
Gender	
Male	68.1
Female	31.9
	100.0

Field Procedures

The survey was conducted in two stages: an initial mail-back survey with minimal follow-up, and an intensive phone survey of a sample of those who did not respond to the mail-back survey. This design was intended to test both the feasibility of using a mail survey with limited follow-up in a population of injured workers, and to assess the nature and magnitude of effect of non-response bias on patient-reported satisfaction and outcomes.

The first stage was designed and fielded as a mail-back self-administered questionnaire. The follow-up effort was deliberately limited to two full mailings and a reminder

postcard, in order to match the conditions of administration that would likely exist in the HCO program, or if the survey were to be routinely administered by managed care organizations, employers, or claims administrators.

The mailings were initially sent to the names and addresses provided by participating organizations. The first mailing, consisting of a cover letter, questionnaire, and stamped return envelope, was sent to the full sample of 802 eligible injured workers on April 18, 1996. Approximately three weeks later, on May 9, 1996, a reminder postcard was sent to workers who had not already returned a completed questionnaire. An additional mailing, again with a cover letter, questionnaire, and stamped return envelope, was sent to all persons who had not yet returned a completed questionnaire on May 31, 1996. The cut-off date for accepting questionnaire returns from the mail-back survey was August 1, 1996. Other than requests for post-office forwarding addresses, no special tracing procedures were used in the mail-back survey; again, this was intended to mirror a level of effort likely to be feasible in routine administration.

The second stage of the field effort consisted of a telephone follow-up of a random sample of persons who had not returned a completed questionnaire by July 15, 1996. We attempted to interview a random sub-sample of 135 eligibles (those who had not returned the self-administered questionnaire) between July 15 and September 12, 1996. In order to achieve a response rate of 60% or higher, we did extensive tracing of this sub-sample, including a) directory assistance, b) post office forwarding address requests, c) on-line Department of Motor Vehicles searches, d) Haines on-line reverse directories, and e) Metronet on-line searches with address inputs and phone number outputs. Follow-up included a large number of callbacks to boost response rate - averaging over seven calls per completed interview. The phone protocol included a "persuasion interview" which asked respondents for reasons they had not responded to the mail-back survey and offered a $5 incentive payment to complete the phone interview.

III. Survey Results

A. Response Rates and Their Correlates

1. Completion Rates in the Mail-Back Survey and the Telephone Follow-Up

Mail-back survey: Table 3 shows the field outcomes for the mail-back survey. Of 802 mailed self-administered questionnaires, we received a total of 238 completed questionnaires by the cutoff date of August 1, 1996. About 44% of the completed surveys were returned after the first mailing. Only four persons (0.5%) sent a note of explicit refusal. Adjusting for language problems (three identified cases) and loss through death (one case), the response rate was 238/798, or 29.8%. Fifty-one (6%) of the eligible respondents were classified as "unable to locate" due to return of the first mailing with no

forwarding address. The remaining 505 persons (63%) presumably received the questionnaire, but failed to return it.

Table 3: **Field outcomes of the mail-out survey**

Field Code	N	Percent
Did not return SAQ	505	63.0%
Returned completed SAQ	238	29.7
Mailing returned incomplete with no forwarding address	51	6.4
Refusal	4	0.5
Known language problem	3	0.4
Deceased	1	0.1
Total	802	100.1

Limited information from the administrative data provided by participating organizations can be used to make comparisons between the mail-back self-administered survey respondents and non-respondents. As Table 4 shows, there were only minor differences between these groups with respect to type of organization from which cases were obtained (self-insured, SCIF, or medical care organization) or broad categories of injury (back strain or sprain, other strain or sprain, other injury). However, there is a pattern of selection bias seen frequently in random samples from the general population - a disproportionately large proportion of respondents are women and older persons [16]. The implications of these demographic differences are examined below.

Table 4: **Comparisons between respondents and non-respondents in the mail-back survey based on information from the administrative database**

Characteristic	Respondent Sample N=238 %	Non-Respondent Sample N=546 %
Type of Organization		
Self-insured	9.2	7.1
State fund	32.4	32.4
Medical organization	58.4	60.5
	100.0	100.0
Nature of Injury		
Back sprain/strain	38.7	32.6
Other sprain/strain	6.3	7.3
Other injury	55.9	60.1
	99.9	100.0
Age		
less than 30	28.2	38.8
31-50 years old	52.5	52.0
51 and over	19.3	9.2
	100.0	100.0
Gender		
Male	60.5	71.3
Female	39.5	28.7
	100.0	100.0

Phone follow-up protocol: The follow-up phone survey was designed to achieve a higher response rate than the mail-out survey, in order to allow assessment of non-response bias with respect to responses to questionnaire items. This sample consisted of 135 persons randomly selected from those persons who did not return a mail-back questionnaire and who were not known to have language difficulties.

Table 5 summarizes the field outcomes for the telephone follow-up survey. Seventy five interviews were successfully completed. Sixteen individuals (12%) had difficulty speaking English, and were coded incomplete "due to language problems". Most of these persons were Spanish-speaking. Only 5% of the follow-up sample refused to be interviewed. We were unable to locate about 4% of the sample. Since the questionnaire was designed for English-speaking respondents, the response rate was thus defined as

completed interviews divided by the sample size minus known language problems, or 75/(135-16), for a response rate of 63%.

Table 5: **Outcomes for the telephone follow-up study**

Field Code	N	Percent
Completed telephone interview	75	55.6
Could not complete interview after several contacts (but no refusal or language problem)	31	23.0
Known language problem	16	11.9
Refusal	7	5.2
Unable to locate	6	4.4
Total	135	100.1

2. Associations between Survey Responses and "Stage of Response"

When the response rate to a self-administered questionnaire survey is low, inferences about the population can be drawn only if we know that there is no significant response bias. In other words, our confidence in extrapolations from the SAQ protocol increases to the extent that SAQ respondents and non-respondents answer important survey questions in substantially similar ways. Large differences in the response distributions of SAQ and phone respondents would suggest that extrapolations from an SAQ with low response rate need be done with caution. The telephone follow-up survey was designed to address this issue. In this section, we examine the similarities and differences between SAQ and phone respondents and try to explain such differences when they occur.

Table 6 shows a classification of respondents by type of data collection (telephone or mail) and, for mail-back respondents, by whether the completed questionnaire was obtained after the first, second, or third mailing. We used this table to construct "stage of response" categories that can be used to make comparisons pertaining to distributions of

responses to questionnaire items. For this purpose, the 489 persons who did not participate in either the mail-back or phone follow-up are regarded as "missing cases" on the stage variable and are thus excluded from the comparisons presented in this section.

Table 6**: Classification of respondents and non-respondents by type of data collection and stage of mail-back survey**

Returned complete SAQ	Respondent in:			
	Mail-back	Phone follow-up	Non-Respondents in mail-back and phone	Total
Returned after first mailing	104	*dna*	*dna*	104
Returned after postcard	31	*dna*	*dna*	31
Returned after second mailing	103	*dna*	*dna*	103
Did not complete SAQ	*dna*	75	489	554
Total	238	75	489	802

Note: *dna* means "does not apply".

Stage of response is defined in three categories. Stage 1 (N=104) consists of persons who returned the SAQ after the first and before the second mailing; Stage 2 (N=134) consists of those who returned the SAQ after the second or third mailing, but before the cutoff date; and Stage 3 (N=75) includes persons who completed a phone interview in the follow-up.

Stage 1 versus Stage 2 represents two levels of difficulty of getting a reply to a mail-back SAQ. Large differences in responses of respondents at these two stages would suggest that "level of pursuit" be considered as a correlate for non-response bias in mail-out studies of this population.

Stage 1 and 2 versus Stage 3 represents both an additional level of difficulty in obtaining a response and difference in mode of data collection (phone versus SAQ). A third factor is the average length of time between injury and return of the SAQ or completion of the

interview, which increases with each stage of response due to the schedule for data collection.

Table 7 shows that the age and gender contrast between mail-back respondents and non-respondents exists also in comparisons of stage of response. Percent male increases with each stage of response, with Stage 3 approximating the percent male found among persons who did not respond to either the mail or telephone survey. Percent of persons over age 51 declines with stage of response, as each succeeding contact tends to obtain larger proportions of younger and middle-aged respondents. To the extent that age and gender are related to other kinds of responses - e.g. to subjective evaluations of well-being and quality of medical care - we might expect to see additional correlates of stage of response in analyzing responses to questionnaire items. This issue is addressed below.

Table 7: **Association between stage of response and age and gender**

Attribute	Stage 1 Return after 1st Mailing N=104	Stage 2 Return after 2nd or 3rd Mailing N=134	Stage 3 Interviewed by Phone N=75	Non-Respondent N=489
Percent Male	54.8	64.9	70.1	71.3
Percent Age 51 and over	25.2	18.2	6.7	9.6

There were no significant differences in distribution of responses to questionnaire items between Stages 1 and 2 (early or late return of the mail-back questionnaire). Therefore, our analysis of the association between stage of response and responses to survey questions emphasizes the mail-back versus phone mode of administration.

We examined the differences and similarities between mail-back and phone respondents with respect to response distributions for each item. Because the phone follow-up sample is relatively small, standard test statistics could detect only rather large differences (i.e. more than 10 percentage points). Thus, wherever differences suggested a relationship, we used a simple goodness-of-fit Chi-squared test of independence with a liberal probability criterion as a filter for further examination. Although this procedure undoubtedly over-emphasized certain differences, it produced a profile that fit well with a simple inspection of the marginal distributions associated with the two protocols.

Two findings are of note. First, the telephone respondents report better perceptions of outcome on many items pertaining to post-injury well-being. Second, there are only a few scattered and unsystematic differences between mail and phone respondents with respect to other questionnaire items.

Table 8 shows differences between respondents with respect to several items related to post-injury well-being; each of these comparisons suggests that the phone respondents perceive a higher level of well-being than the mail respondents. More mail respondents rate their health now as worse than before the injury. Almost two-thirds of the mail respondents - vs. half the phone respondents - regard the injury as having some or a major effect on their lives; although most respondents have no difficulty with daily chores, substantially more SAQ respondents have difficulty with each of these activities. More phone than SAQ respondents regard themselves as fully-recovered. More than 40% of the mail respondents report pain due to the injury "almost every day" or "all the time", compared with 23% of the phone respondents. Over 25% of phone respondents report difficulty making mistakes due to emotional problems, compared to 44% of the SAQ respondents; 12% of SAQ respondents versus no phone respondents mention that their injuries involved stress. Finally, a quarter of the SAQ respondents had not returned to work at the time they completed the questionnaire, versus just 13% of phone respondents - an interesting finding even though a slightly higher proportion of phone respondents indicated they had never missed work.

Table 8**: A profile of differences between mail-back SAQ respondents and phone respondents**

Questionnaire Item	Mail-Back Respondents N=238 %	Phone Respondents N=75 %
Injury involve emotional or mental stress?(Q3B10)		
Yes	11.6%	0.0%
No	88.4	100.0
	100.0	100.0
How is health now compared to before injury? (Q4)		
Better	17.1	21.6
Same	35.9	45.9
Worse	47.0	32.4
	100.0	100.0

Questionnaire Item	Mail-Back Respondents N=238 %	Phone Respondents N=75 %
How much does injury affect life today?(Q5)		
Some to big effect	64.7	49.3
No effect to little effect	35.3	50.7
	100.0	100.0
Difficulty lifting groceries, last 4 weeks? (Q7A)		
No difficulty	45.4	73.0
A little difficulty	17.9	2.7
Some difficulty	22.0	18.9
A lot of difficulty	14.7	5.4
	100.0	100.0
Difficulty climbing stairs, last 4 weeks? (Q7B)		
No difficulty	52.4	76.7
A little difficulty	16.3	12.3
Some difficulty	20.7	8.2
A lot of difficulty	10.6	2.7
	100.0	100.0
Difficulty opening jars, last 4 weeks? (Q7C)		
No difficulty	63.8	77.3
A little difficulty	14.8	5.3
Some difficulty	12.9	10.7
A lot of difficulty	8.6	6.7
	100.0	100.0
Pain due to injury in last 4 weeks? (Q.8)		
Not at all	26.3	39.2

Once in a while to several times a week	30.6	28.4
Almost every day to all the time	42.7	23.0
	100.0	100.0
Difficulty making mistakes due to emotional problems in last 4 weeks? (Q10B)		
No difficulty	55.9	73.6
Little to some difficulty	35.9	25.0
A lot of difficulty	8.2	1.4
	100.0	100.0
Worked for pay since injury? (Q26A)		
Never missed work	21.5	26.7
Returned to work after being off	53.2	60.0
Have not returned to work	25.3	13.3
	100.0	100.0

We considered four explanations for these differences. First, there is evidence that age and gender are selective factors in the determination of early versus late response. It is possible that older persons and women, who are more likely than younger persons and men to respond via mail in early contacts, are also more likely to report poorer outcomes.

A second hypotheses is that persons who respond to the mail SAQ are different in other ways which are associated with poorer self-reported outcome; for example, mail response might be correlated with greater severity of injury, which could also correlate with poorer perceived well-being after injury.

Thirdly, it may be that answering questions in the SAQ produces more complaints about well-being than answering the same questions by phone, in which case the differences would be attributable to mode of data collection.

Finally, it is possible that the differences in perceived well-being after injury result from a correlation between stage of response (i.e. mail SAQ vs. phone) and the length of time between occurrence of injury and collection of data; this length of time is an indirect measure of "recovery" time. As noted earlier, the correlation between recovery time and mode of data collection is a direct consequence of the pilot study design, in which the follow-up phone sample is a subset of the mail SAQ non-respondents.

Although these explanations are inter-related and thus cannot be tested independently, a partial test of their comparative explanatory power can be accomplished by the use of multivariate analysis. We dichotomized each of the ten responses examined in Table 8, using logistic regression to examine the relation between these dichotomized responses and four predictor variables: male versus female, age (less than 30, 31-49, 50 and over), SAQ vs. phone interview, and elapsed months from self-reported date of injury to date of interview or date of return of the questionnaire.

Controlling for age, gender, and elapsed time from injury to data collection SAQ/mail-back versus phone interview adds to the prediction of four out of ten of the responses in Table 8 (poorer health relative to pre-injury, trouble with lifting and climbing stairs, and more frequent pain). For the other six responses in table 8, differences in mode of data collection accounted for none of the differences when controlling for the other factors. Although selection bias with respect to age and gender appears to contribute slightly to the variation in well-being responses, it does not account for the pattern of differences observed in Table 8. Time from injury to data collection is a dominant predictor in predicting return to work, which increases with elapsed time even after controlling for other variables.

These mixed results suggest that, while age, gender, and "recovery time" may contribute to the contrasts we observe in Table 8, they cannot completely explain them. Some of these differences may be due to a) differences in reporting on the SAQ versus the phone interview or b) other factors (e.g. perceived severity of injury) that may be related to a tendency to respond early but not strongly correlated with elapsed time from injury to data collection. Differences in reporting could be related to questionnaire wording; as discussed below, an analysis of item non-response identified some problems with the wording of questions related to function in everyday activities. Furthermore, the failure of "recovery time" to explain more of the differences found in Table 8 could be in part due to measurement problems with reported dates of injury, also discussed below.

With the exception of the functional outcomes discussed above, there are few associations between the mail vs. phone responses. SAQ respondents did report more discussions with doctors about work restrictions and the job.

Of particular importance is that we observed no significant differences with respect to most of the subjective ratings of the quality of medical care for injuries, as shown in Table 9.

Table 9**: Comparisons between mail-back SAQ respondents and phone interview respondents with respect to respondent ratings of the quality of medical care for injury**

Type of Rating	Mail-Back SAQ Respondents N= 238	Phone Interview Respondents N=75
Percent reporting they were very or somewhat satisfied with health care for injury (Q16)	74.5	70.7
Percent reporting no trouble getting medical care when first injured (Q13)	67.6	73.3
Percent who report that doctor suggested job changes to help injury heal (Q17B)	68.8	67.6
Percent who report they were told how to avoid re-injury at work (Q17C)	55.7	56.2
Percent who report they were told about work restrictions or job changes needed or that none were needed (Q17D)	81.0	68.0
Percent reporting doctor talked some or a lot about job and work (Q19B)	69.0	55.4
Percent reporting doctor understood kind of work very well or fairly well (Q19C)	79.2	73.4
Percent reporting doctor understood how injury happened well or fairly well (Q19D)	88.1	89.3
Percent reporting doctor providing most care was excellent or very good listening to patient (Q20B1)	87.3	86.3
Percent reporting doctor providing most care was excellent or very good showing courtesy (Q20B2)	72.4	68.0

Table 9**: Comparisons between mail-back SAQ respondents and phone interview respondents with respect to respondent ratings of the quality of medical care for injury**

Percent reporting doctor providing most care was excellent or very good explaining things (Q20B3)	69.1	69.4
Percent reporting doctor providing most care was excellent or very good in thoroughness of exam (Q20B4)	61.6	60.0
Percent reporting doctor providing most care was excellent or very good in figuring out what was wrong (Q20B5)	63.1	70.7
Percent reporting doctor providing most care was excellent or very good understanding how injury would affect job (Q20B6)	56.4	62.0

3. Item non-response

The amount of non-response to a particular question can be interpreted as a measure of a) lack of clarity in posing the question, b) the difficulty of making a response (e.g. questions that invoke recall memory), c) the respondents' reluctance to disclose information about the item, or d) some combination of these. The study protocol did not provide for additional contacts to reduce item non-response after return of the mail SAQ.

We examined item non-response among mail and phone respondents. Although the phone follow-up sample is too small to provide stable estimates of item non-response for phone interviews, the expected lower rates for an interviewer-administered protocol versus a self-administered questionnaire were generally borne out. Some of the problems of item non-response for the SAQ appear to have been mitigated by the presence of an interviewer.

There were approximately 90 questions on the SAQ. The item non-response rate for four-fifths of them was 2% or less. Only nine items had non-response rates of 5% or more, and nearly all of these "excess" rates can be explained by the factors enumerated above.

The questions generating the most item non-response dealt with recent difficulties with daily tasks (Questions 7A-C on p. 3 of SAQ) and recent difficulties with activities because of emotional problems (10A-B p. 4). Non-response to these five questions ranged from 8% to 18%. In this case, the problem appears to have been caused by the way response categories were worded and placed on the printed page. For example, question 7A asks "During the last four weeks, how much difficulty have you had with each of the tasks listed below? A. lifting or carrying a full bag of groceries or something else that weighs about 10 pounds." From left to right, the response categories are "No difficulty in the last 4 weeks OR NOT caused by injury", "A lot of difficulty", "some difficulty", and "a little difficulty". The first category attempts to serve two functions - to screen out persons with no difficulty, and persons whose difficulty was not due to their work injury.

Hindsight suggests that it would have been better to separate these questions (e.g. was the activity difficult? if so, was the difficulty due to the injury?). Additionally, the response categories were not ordered by level of difficulty; the visual break separating "No difficulty" from "A lot of difficulty" was evidently not enough to prevent confusion due to the violation of left- to- right order. This too could be remedied by asking two questions for each question instead of one.

All remaining instances of significant item non-response involved questions that ask for dates, times, or income. Questions 2 (date of injury), 14 (amount of time from injury to first seeing a doctor), and 26B (date of first return to work) each had about a 5% non-response. In these cases, recall failure may induce non-response as the respondent strives to avoid making an inaccurate response. The similar level of non-response for the income item (40) is typical in survey work, and may reflect privacy considerations.[16]

III. B. Data Quality and Measurement Issues

1. Validation of Self-Reports of Injury Type

Comparison of self-report of type of injury and limited administrative data validates the self-report information as a crude measure of nature of work injury. Table 10 shows the association between two self-reports of injury type (Q3A1 body part, and Q3B1 nature of injury) and a three-category injury classification as constructed from administrative data supplied by claims administrators and medical organizations. Of those injuries classified administratively as "back strain or sprain", about 62% are self-reported as back or neck sprain, and another 24% as sprain, strain or other injury to joint or muscle. This represents about 86% agreement in gross categorization of back strain/sprain. About half of the "other strains" and "other injury" categories in administrative data appear to be consistent with the self-reports. In the absence of medical records, the source of disagreement remains unknown. However, given the crude categories used in the survey,

and the likelihood of misclassification or reporting error in both administrative and self-report, the amount of disagreement is not surprising.

Table 10: **Administrative data versus self-report pertaining to nature of injury: percent distribution of self-report by administrative injury category, N=308**

	Self-report, back or neck strain or sprain	Self-report, other back or neck injury	Self-report, sprain or strain of other body part	Self -report other injury	Totals
Administrative data					
Back sprain	62.3	7.0	23.7	7.0	100.0%
Other sprain	19.0	9.5	47.6	23.8	100.0%
Other injury	16.2	3.5	28.9	51.4	100.0%

2. Variation in Subjective Evaluations of Medical Care

An important aspect of data quality in subjective ratings is the degree to which survey responses are concentrated into a few of the several categories offered to the respondent. In ratings of the quality of medical care, it is known that the response distributions tend to be skewed in the direction of positive evaluations [17]. This may represent that people are truly happy with the care they receive, or a tendency to report favorably regardless of actual experiences, or a bias toward positive response introduced in the way the questions are asked.

We assessed whether there is sufficient variation in responses to perform routine statistical analyses related to the sources and consequences of patient satisfaction, and whether there are sufficient negative responses to attempt to profile characteristics of negative responders and compare them with positive responders.

For this analysis, we pooled the mail and phone responses, examining ten subjective evaluations among 313 respondents. Items included one item on general satisfaction with care (Q16), four items related to patient perceptions of occupational medicine (Q19B-D and 20B6), and five questions about doctor-patient communication and exam thoroughness. Three measures of the distributions of responses are shown in Table 11: a general index of variability, the ratio of the frequencies of the most to least favorable response, and the prevalence of responses that can be regarded as “negative”.

Table 11: **Measures of variation and skewness in the distributions of subjective ratings**

Question/# of response categories	Type of Response Categories	Index of Response Variation (min=0, max=1)	Ratio of Most Favorable to Least Favorable Response	Prevalence of Unfavorable Responses (2 least favorable)
Q16, General satisfaction/4	Very dissatisfied to very satisfied	.68		26%
Q19B, Talk about job/4	A lot to not at all	.79		34%
Q19C. Understand Job/4	Very well to not at all	.63		22%
Q19D, Understand how injury happened/4	Very well to not at all	.49		12%
Q20B1, Listen/5	Excellent to poor	.66		13%
Q20B2, Courtesy/5	Excellent to poor	.61		9%
Q20B3, Explain/5	Excellent to poor	.64		12%
Q20B4, Exam/5	Excellent to poor	.69		15%
Q20B5, Diagnosis/5	Excellent to poor	.69		16%
Q20B6, Understand effect of injury on job/5	Excellent to poor	.77		24%

As expected, the most favorable responses are consistently more numerous than the least favorable. Nevertheless, the variance indicator, which ranges from 0 (all responses fall in one category) to 1 (each of the possible responses are equally likely), shows that there is sufficient variation in responses for statistical analysis, even though much of this variation occurs among the favorable responses. The prevalence of unfavorable responses, however, ranges from a low near 10% to a high of over 30%; of particular interest is that more than one-quarter of injured workers express some level of generalized dissatisfaction with the health care they received for their injury. There are evidently a sufficient number of discontented patients to conduct an analysis of their characteristics, assuming a sample at least as large as the one obtained in the pilot survey.

3. Dimensionality of Responses Pertaining to Occupational Medicine

We developed several questions specifically related to occupational medicine, and examined the inter-correlations among responses to these occupational medicine questions to determine if these items reflect a single dimension of care, as seen from the patient's perspective. Four occupational medicine questions asked the patient to report on what the providers did - did they suggest changes in the job to help the injury heal?, tell how to avoid re-injury?, suggest work restrictions or say none were needed?, or talk about the patient's job (Q17B-D,19B). Three items asked for the patient's evaluation of how well the primary treating physician understood the patient's job, how the injury would affect the job, or how the injury occurred (Q19C,D,20B6). All response categories were coded so that the high scores represented poor evaluations or an action not taken by the providers.

Although there are positive correlations between most of the indicators, a factor analysis summarized in Table 12 suggests two separate dimensions in the occupational medicine item. Factor I represents an evaluative component pertaining to how well the doctor understands. Factor II pertains to whether the patient was told much or little. The internal coherence of these factors, and their mutual independence, may result from both differences in the item structure (evaluation vs. report of behavior) and methodological factors such as placement in the questionnaire and response categories.

The factor analysis of occupational medicine items suggests that patients may differentiate their evaluation of a provider's ability to understand the job-related aspects of an injury, versus their actions related to this understanding. A physician may understand but not tell, tell but not understand, tell and understand, or neither tell nor understand. Given this dimensionality, it may not be warranted to construct a single index pertaining to occupational medicine based on these items.

Table 12**: Results of factor analysis of occupational medicine questions**

Question	Loading on Factor I	Loading on Factor II
Q17B: Drs. suggest job changes to help injury	.106	.794
Q17C: Told how to avoid re-injury at work	.251	.653
Q17D: Told about changes necessary to continue work	-.020	.576
Q19B: Amt Dr. talked about R's job and what R does at work	.749	.226
Q19C: How well Dr. understood things R does on job	.889	.022
Q19D: How well Dr. understood how injury happened	.770	.104
Q20B6: How well Dr. understood injury's effect on ability to do job	.767	.084
Percent of variance explained by factor	41%	18%

4. Agreement Among Alternative Measures of Access to Care

There are three measures of access to care for injuries in the questionnaire: an overall rating of trouble getting care for the injury (Q13), how soon after injury care was received (Q14A), and waiting time in office or exam room (Q14C). The overall evaluation was generally in accord with the more specific indications of time to care. For example, about 65% of those reporting "no trouble" in getting care measured time from injury to care in hours, whereas only 38% of those reporting "a lot of trouble" measure time to care in hours. Although we have no independent measure of waiting times to use for validation, these results suggest that the overall evaluation of access is constructed from the raw material of actual waiting times.

5. Correlations between general and specific ratings

Table 13 examines correlations between the general rating of satisfaction with medical care and several specific evaluations and self-reports. There are significant correlations between general ratings and assessments of the quality of the doctor-patient interaction (e.g. listening, showing respect, explaining); these correlations are consistent with patient satisfaction findings in other arenas. For the subset of workers who report pain associated with their injuries and who have recently seen a physician, overall satisfaction with care is strongly related to ratings of how well the doctor helped in pain management (Q9). On the other hand, trouble with access and reports about what doctors told patients had a negative correlation with general satisfaction.

Table 13: **Correlations between general satisfaction and selected reports and subjective ratings, average N=300**

Correlate of General Satisfaction(Q16/4 response categories)	Correlation Coefficient
Trouble had Get Med Care when First Inj (Q13/4 response cats.)	-.260
Drs Suggest Job Changes to Help Injury (Q17B/2 response cats.)	.160
Told How to Avoid Re-Injury at Work (Q17C/2 response cats.)	.269
Told About Changes Nec to Continue Work (Q17D/2 response cats.)	.018
Amt Dr Talk abt R's Job & what do at Wrk (Q19B/4 response cats.)	.211
How Well Dr Undstd Things R does on Job (Q19C/4 response categories)	.311
How Well Dr Undrstd how Injury Happened (Q19D/4 response cats.)	.439
How Good Job He/She Did Listening to R (Q20B1/5 response cats.)	.528
The Courtesy and Respect He/She Showed (Q20B2/5 response cats.)	.468
How Good Job Explaining Things (Q20B3/5 response cats.)	.481
How Thorough/Careful Exams & Treatment were (Q20B4/5 response cats.)	.511

Ability to Figure Out what was Wrong (Q20B5/5 response cats.)	.568
How Well Understood Injury Affect Ability to do Job (Q20B6/5 response cats.)	.525
How helpful doctors/health professionals have been in managing pain in last four weeks (Q9/4 response cats.) N= 119 Persons with pain who've been to doctor in last 4 weeks	.664

6. Agreement between self-reported date of injury and date of injury in claims file

299 respondents gave a codable date when asked to report the month and year of their injury (Q2). The reported date of injury was compared to the date of injury in the administrative database from which the survey sample was drawn, as reported by the claims administrator or medical care provider. Of the respondents with both dates recorded, about 77% gave dates of injury within one month of the claim date. On the other hand, 14% of the respondents gave dates *later* by more than a month, and 9% gave dates *earlier* by more than one month. of the latter, 10 respondents gave dates that precede the claims date by more than 12 months.

The substantial number of differences are large enough to make random recoding errors or lapses of memory unlikely explanations. While most respondents are likely reporting about care for the same injury that made them eligible for the survey, there appears to be a subset of respondents who are answering the survey about either a) an earlier injury associated with a re-injury that generated the claim making them eligible for the survey or b) a re-injury related to the earlier injury that produced the claim. Differences in self-reported and administrative date of injury were more extreme for the SAQ/mail respondents than for the phone respondents, and overall variance was larger for the SAQ.

III. C. Analysis of subjective evaluations of medical care

To what extent can patient ratings be used as indicators of the nature and quality of care for work injuries? When are negative ratings a "red flag" and positive ratings a genuine indication that proper care is being provided? How much are such judgments influenced by factors that are unrelated to the adequacy of medical care? To shed light on these questions, it is important to investigate the patterns of association between subjective reports and other factors that may influence them in concert with, or apart from, the quality of the care received. Although extensive analysis of this type is beyond the scope of this project, we examined the relation between reports of overall satisfaction with care and selected potential "confounders".

We chose three classes of variables as possible predictors of general satisfaction: 1) demographic and socioeconomic factors consisting of age, gender, educational level, and marital status; 2) factors related to self-reported stage of recovery from injury; and 3) patient ratings and reports pertaining to specific aspects of care. Of these, only the latter two proved to be plausible explanatory factors. Age, gender, educational attainment, and marital status were not associated with overall ratings of satisfaction with care in our sample; crude measures of type of injury and return to work were also not related to general satisfaction.

Most self-reported indicators of stage of recovery from work injuries were strongly related to satisfaction with care; in general, the greater the sense of recovery, the greater the satisfaction. Dissatisfied ratings tended to be concentrated among patients with frequent pain, marked disabilities in everyday activities, perception of little improvement since the injury, and a belief that the injury still has a big effect on their lives.

The correlations in Table 13 suggest overall satisfaction is also related to evaluations of specific aspects of treatment for the work injury. Reexamination of these associations using other techniques (measures of association for cross-classifications and tests based on Chi-squared statistics) confirmed these results. In particular, the cluster of evaluations pertaining to doctor-patient interactions (Q20B1,2,3), thoroughness of examination (Q20B4), ability of physician to diagnose and treat (Q20B5), and how well the doctor understood the impact of the injury on the job (Q20B6) were all strongly and significantly associated with global satisfaction with medical care.

Additional analysis of general satisfaction using multivariate methods is warranted in larger samples, to further understand the composition of the satisfaction ratings - what factors appear to predict satisfaction best? are ratings composed of many small influences or a few dominant factors? to what degree are ratings of health care contaminated by factors external to the quality of care? Table 15 shows the result of one illustrative trial-and-error effort to build a predictive model using logistic regression analysis to analyze the responses to the overall satisfaction question, expressed as a dichotomous “very or somewhat dissatisfied” versus “very or somewhat satisfied” response. The set of predictors used for this analysis was chosen from among the experience ratings and recovery variables after a non-exhaustive search for a parsimonious prediction equation. By eliminating redundant and ineffective predictors, we arrived at an equation that represents general satisfaction as a function of outcomes perceived by the patient (e.g. pain and sense of recovery) and physician ability to diagnose and treat, again as perceived by the patient. (Table 14 shows how responses were coded for this analysis).

Table 14: **Coding of variables for logistic regression analysis of general satisfaction ratings**

	Value	Freq	Parameter Coding (1)	(2)	(3)	(4)
Q16-Overall Sat.						
Dissatisfied	1	79				
Satisfied	0	222				
Q8A-Pain						
NtAt All	0	90	1.000	.000	.000	.000
All Time	1	44	.000	1.000	.000	.000
Not Cnst	2	69	.000	.000	1.000	.000
SvrlX/Wk	3	34	.000	.000	.000	1.000
OnceWhil	4	64	.000	.000	.000	.000
Q20B5-Ability to Diagnose and Treat						
Excl	1	99	1.000	.000	.000	.000
VeryGood	2	97	.000	1.000	.000	.000
Good	3	58	.000	.000	1.000	.000
Fair	4	24	.000	.000	.000	1.000
Poor	5	23	.000	.000	.000	.000
Q6-Perceived Stage of Recovery						
FullRcov	1	94	1.000	.000		
RcovSome	2	173	.000	1.000		
No Impv	3	34	.000	.000		

Table 15: **Results of logistic regression of general satisfaction with care received for injury** (Q16, coded as 1=dissatisfied, 0=satisfied)

Each of these factors is related to the dependent variable - dissatisfied responses - in the expected direction. Patients who report pain and feel they are not recovered tend toward negative ratings. Patient who fell that the physicians had poor abilities to "figure out what was wrong and what needed to be done" were naturally inclined to give poor overall evaluations. Combining these factors, the fitted logistic equation is able to predict the observed rating about 83% of the time. A measure analogous to the squared multiple correlation of ordinary regression analysis shows that approximately 34% of the total variation in the rating is explained by this logistic regression model.

It is worth noting, however, that the prediction equation fails to predict about half (30 of 79) of the poor ratings. Adding other indicators which measure different facets of the evaluation process may improve prediction to some extent. Table 16 show, for example, that respondent ratings of pre-injury job satisfaction and having a good relationship with the supervisor are associated with favorable ratings of overall satisfaction with medical care. Consulting a lawyer is associated with greater dissatisfaction, while union membership had no relationship to ratings of medical care.

Q6= Subjective rating of level of recovery
Q20B5= Patient rating of doctor's ability to diagnose and treat
Q8A=Patient report of pain felt in last 4 weeks

Variable	B	S.E.	Wald	df	Sig	R
Q6(Recovery)			8.5371	2	.0140	.1144
Q6(1)	-1.9643	.8523	5.3123	1	.0212	-.0978
Q6(2)	-1.5940	.5603	8.0930	1	.0044	-.1326
Q20B5(Diagnosis and Treatment)			41.8665	4	.0000	.3126
Q20B5(1)	-5.2140	1.1352	21.0974	1	.0000	-.2348
Q20B5(2)	-5.1339	1.1272	20.7450	1	.0000	-.2326
Q20B5(3)	-3.9776	1.1212	12.5862	1	.0004	-.1748
Q20B5(4)	-2.6367	1.1720	5.0610	1	.0245	-.0940
Q8A (Pain)			7.9776	4	.0924	.0000
Q8A(1)	-.9311	.7432	1.5695	1	.2103	.0000
Q8A(2)	.0623	.6413	.0094	1	.9227	.0000
Q8A(3)	.6667	.5203	1.6416	1	.2001	.0000
Q8A(4)	1.0282	.5728	3.2222	1	.0726	.0594
Constant	4.4869	1.2804	12.2794	1	.0005	

Table 16: **Associations between general satisfaction with medical care for injury and selected additional factors**

Factor	Percent "very satisfied" or "somewhat satisfied" with health care received for injury (Q16)	
	%	N
1. Satisfaction with pre-injury job (Q24)		
Very satisfied	76.5	168
Somewhat satisfied	70.9	103
Somewhat dissatisfied	73.3	30
Very dissatisfied	60.0	10
2. Relationship with immediate supervisor at time of injury (Q25)		
Excellent	76.6	142
Good	77.6	117
Fair	58.6	29
Poor	56.6	23
3. Currently a member of a labor union (Q31)		
Yes	73.5	98
No	73.6	214
4. Talk with/hire attorney to represent for claim (Q32A&B)		
Did not talk to attorney	77.1	245
Talked to attorney/didn't hire	52.6	19
Hired attorney	63.8	47

III.D. Return to Work

Return to work outcomes were assessed in a series of questions (Q26-30) near the end of the questionnaire, as shown in Table 17. Nearly four-fifths of respondents reported that they had returned to work at the time of data collection. 23% said they had never missed work, 55% had been off work due to injury but had returned to work, and 22% reported that they had not yet returned to work.

Using elapsed time from self-reported date of injury to self-reported first return to work, about 70% of those who returned to work after an initial period of lost work time were back to work within about two months; about 90% were back to work within four months. Nearly all of those who returned to work (93%) went back to the same employer, and most had either no subsequent lost work time (55%) or less than 10 days subsequent lost time (24%) by the time of data collection. About 30% of the total sample reported that they had not worked at all in the four weeks prior to filling out the questionnaire.

Among those who returned to work after their injury, 38% reported that their work had been modified or that their hours had been restricted due to the injury. The majority of these respondents (63%) were satisfied with the job changes. Of all respondents who had worked in the last four weeks, 31% had some difficulty in performing work, 50% experienced limitations in the kind of work they could do, and 20% had cut down their work hours.

Among persons who say they are “fully recovered”, nearly all have worked since injury; 72% of those who had “recovered some” had worked since injury, versus only 47% of those reporting “no improvement”. It is reasonable to expect that, given that return to work and satisfaction are both correlated with recovery, they would themselves be correlated. However, we found no detectable association between the overall measure of return to work (work for pay since injury) and general satisfaction with medical care.

Table 17: **Self-reported return to work outcomes**

Return to Work Outcome	Percent of Respondents Reporting	
	N	%
Return to Work		
Never missed work	71	23
Missed work, returned to work	171	55
Not returned to work	70	22
Lost -time since Return to Work		
No subsequent lost time	93	55
Less than 10 days	52	24
More than 10 days	35	21
Current Work Status		
Worked in last four weeks	42	70
No work in last four weeks	23	30

Discussion and conclusions

Survey administration: The construction of a sample for assessing patient-reported outcomes and satisfaction depends on at least three considerations besides sample size: a) the definition of the target population of workers (e.g. workers with at least three days of temporary disability or with particular injuries or a particular dollar amount for medical only cases--- > $2,000.); b) the range of dates of injury to be included; and c) the amount and accuracy of information available from the records used to construct the sample. The last of these is of particular importance, since for both mail and phone survey protocols, costs and response rates vary directly with the accuracy and completeness of names, current mailing addresses, and current telephone contact information.

Several other factors determine response rates in mail-back surveys. Among the most important of these are a) the social and demographic composition of the target population; b) the questionnaire content; and c) the level of effort made to obtain completed questionnaires. Typically, highest response rates can be attained with older females, and in populations with greater education and higher income. In general, the target population of injured workers is biased toward persons who are difficult to recruit into surveys. It is disproportionately young, male, and in occupations requiring less education and characterized by lower wage rates.

Given the limited follow-up, the nature of the target population, and the potential impact of language difficulties (not measured but surely present), SRC researchers felt that a return rate in the neighborhood of 30% is not unusually low.

The higher response rate obtained in the telephone survey was a planned outcome. Phone protocols generally yield better response rates than mail-back protocols[18]; money incentives also help to increase response rates in both data collection modalities[19].

However, the level of effort (and cost) required to obtain a response rate greater than 60% was high, including tracing and multiple calls (averaging 7.3 per completed interview) as noted earlier. Those eligible for phone interviews included only injured workers who were non-respondents in the mail-out survey, suggesting that they were generally more resistant to participation than the SAQ respondents. It is thus likely that the response rate obtained in the telephone survey here is lower than that which would be achieved if the phone survey was the initial data collection effort. Telephone administration of the survey would, in general, yield a higher response rate than the mailed SAQ.

Phone respondents were asked why they had not responded to the mailed SAQ. The typical responses had to do with "being too busy" or "forgot"; almost no respondents suggested that it was because of concerns about filling in the SAQ, or unwillingness to participate.

Non-response bias: The relatively low response rates associated with the mail out protocol naturally provoke concerns about non-response bias. This study suggests that such biases are, in fact, not severe. We measured several differences between mail SAQ respondents and phone respondents. The most pronounced of these differences pertain to responses on several (but not all) questions about post-injury well being. Relative to the phone interview, the mail-back SAQ protocol tends to produce somewhat more frequent reports of problems with post-injury health and well-being.

These reports are not due solely to age/gender selection bias, or solely to recovery bias (i.e. longer elapsed time from injury to interview in phone respondents). It is possible that there were unmeasured differences in severity of injury between SAQ and phone respondents; for example, a slightly higher proportion of phone respondents had never missed work. Another possibility is that a phone interview simply elicits more positive responses about outcome than a self-administered instrument. Or, these differences may reflect a greater willingness of those with self-perceived poor outcomes to respond to a mail SAQ survey.

Whatever the reason, it appears prudent to expect that results using this type of SAQ protocol may be slightly biased toward poorer self-report regarding some recovery outcomes, in comparison with the distributions of recovery outcomes in phone surveys

with higher response rates. Further research on this issue in larger populations of injured workers would be helfpul.

However, there appears to be no such bias with regard to reported satisfaction with medical care. We thus conclude that, in the face of resource constraints which make telephone administration difficult, a mail-out SAQ is a useful tool for obtaining information about satisfaction and self-perceived outcomes from a population of injured workers. Comparisons of self-perceived outcomes in groups of injured workers should take into account any significant differences in the response rates among the different groups. Periodic supplementary phone interviews with non-respondents to assess bias in reports of recovery may be advisable if resources permit.

Timing of data collection with respect to injury: The positive association we observed between direct markers of recovery from injury (e.g. self-assessed recovery, pain, functional incapacity) and satisfaction with medical care has implications for both design of data collection and analysis of results. To make subjective ratings comparable without imposing a requirement for statistical adjustment, it is desirable to standardize the interval between injury and data collection as much as possible. Holding the interval constant also standardizes the recall involved in answering questions about past events. Selection of the interval from date of injury to data collection must balance several factors, including time for recovery, increased difficulty in contact as time from injury increases, and recall bias issues. To the extent that time since injury and recovery are associated, data collection very soon after injury might show lower satisfaction rates than later data collection.

Our intuitive conclusion after this study is that an interval of 6 to 8 months after injury may be optimal for assessment of patient-reported outcomes and satisfaction, although this study did not empirically study this question. Several additional methodological issues are raised by the recovery effect, however, which we have not addressed in this pilot. Because nature of injury is also a determinant of recovery time, it appears that it would be prudent to do a case-mix adjustment or diagnosis-specific analysis if patient surveys are used to compare satisfaction and outcomes among different groups of injured workers or among health care providers.

Focusing on the right injury: Our comparison between self-reported date of injury and claims file date of injury suggests that a sizable minority of respondents anchored their responses at a point in time different from the one we assumed they would use. The pilot did not allow for further investigation to determine whether these differences in time reflected responses based on different injuries, or simply recurrences of the same injury as that for which they were selected into the sample. This finding has implications for date sensitive measures of outcome (e.g. time to first return to work), and for analysis of outcomes which may be subject to a recovery effect.

The SAQ protocol allowed for more divergence between dates than the phone interview protocol. It is possible that the date of injury from the claims record could be incorporated into a mail SAQ instrument, in an effort to focus the respondent in the same time frame as the survey investigator.

Reliability and validity of self-reports in general: With the exception of the problems of date reporting noted above and some special item-non-response issues addressed earlier, we believe the self-reports are reliable and valid indications about patient perceptions of medical care. This assessment is based on our data quality evaluations, and on the structure of the overall subjective ratings of satisfaction with care. We can thus recommend the use of this questionnaire (with the caveats noted above) in the monitoring of injured worker satisfaction with medical care, and self-reported health and functional outcomes after work injury.

Analysis of patient satisfaction and outcomes: Satisfaction with medical care and perceptions of outcome after an injury have multiple determinants and there is substantial variation in both. Because the focus of this study was the validation of the survey instrument, we performed only a limited analysis of the determinants of satisfaction, or the findings with respect to patient-reported outcomes. This introductory exploration suggests that general satisfaction is rooted in the experience of care, and its relation to recovery: freedom from pain, functional ability, and a positive view of interactions with health professionals. Other factors not directly related to the quality of medical care (e.g. relationship with supervisor, attorney representation) may also be related to satisfaction (see Table 16), although these were not statistically significant in this study. Administration of the survey in larger populations of injured workers will doubtless yield important information about the experience of injured workers after injury.

Non-English speaking patients: Approximately 10 % of our sample could not participate in the survey due to language difficulties. Translating questionnaires and interview protocols into Spanish will likely yield benefits to completion rates that justify the costs, particularly in surveys of injured workers with larger sample sizes that those targeted in this project.

The value of information about what injured workers think about their medical care: Injured workers have a unique perspective on the medical care they receive, and on their recovery after injury. Insight into workers' perspective can provide important information for health care providers and organizations, employers, and claims administrators. Such information can point out areas for self-improvement, or be useful in negotiating contracts for services.

For the purchasers of workers' compensation medical care, patient satisfaction surveys are a key component in efforts to assess the value of health care services. Patients can provide important feedback regarding the extent to which providers are gaining the trust

of workers, and communicating in a way that facilitates compliance with treatment regimens. Patients alone can report on the extent to which providers are engaging in discussion about work restrictions and prevention of re-injury.

For employers, workers' perceptions of recovery outcomes after injury and adequacy of may provide important insights into issues that affect workplace morale and productivity. For claims administrators, satisfaction with medical care may impact control over medical treatment, and litigation rates.

Managed care organizations can use patient-derived information to identify areas in which organizational resources may augment physician services (e.g. enhanced case management services), or where more aggressive physician education may be warranted.

Recommendations

The purpose of this study was to develop and test a mail-out protocol for a self-administered questionnaire to evaluate injured workers' satisfaction with medical care and self-perceived functional and health outcomes. We conclude that use of a mail SAQ is useful for this purpose. In this section, we provide recommendations, based on our experience, for the administration of an SAQ to injured workers[2]. A final recommended survey instrument, reflecting revisions made pursuant to this pilot study, is provided as Appendix D.

How many injured workers should be included in a survey?

The survey results will not be reliable if there are not adequate numbers of injured worker respondents. Although there is no definitive number, we recommend that there be at least 200 respondents, or 200 respondents per group (if the survey is being used to compare groups). Assuming a 30% response rate, this would mean 600 injured workers in a sample.

Which workers should be surveyed?

The survey should be administered to injured workers who are likely to have experienced more than casual contact with the workers' compensation medical care system. Currently, we recommend the following inclusion criteria:

a. injured workers who received temporary disability payments or missed at least three days of work

[2] For a more complete description of survey methodology (in a non-workers' compensation population) please consult the HEDIS 3.0 (Jan.1997) Member Satisfaction Survey, Protocols for Sampling and Data Collection available from the National Committee on Quality Assurance, Washington, DC.

b. injured workers who received medical treatment costing more than $2,000.

What information is required to contact the injured workers?

It is extremely important that the contact information for the injured workers be current and accurate. Without good contact information, injured workers cannot be reached to participate in the survey. The following information is required:

a. date of injury
b. name
c. address (use most recent address information)
d. phone number (particularly if any phone follow-up is anticipated)

Who should administer the survey? The survey may be administered by any organization concerned about what injured workers think about their medical care. However, it is likely that the response rates will be higher (and the results thus more reliable) if injured workers are not fearful that their survey results will be seen by anyone who may influence their care, work, or benefits. Thus, we strongly recommend that employers, claims administrators, and managed care organizations consider contracting for survey administration to a neutral survey vendor or survey research organization.

It is imperative that individual responses to the survey be kept confidential. Reputable survey organizations can help to guarantee such confidentiality.

How is the survey administered?

The survey should be mailed to the injured worker 5 and 1/2 months after the date of injury. A reminder post-card can be sent to the worker about 2 weeks after the initial mailing. Another survey should be mailed to those workers who have not responded after an additional 2 weeks. A telephone follow-up to non-respondents can be used to increase response rate; if not feasible to do this routinely, it is recommended that a telephone follow-up to non-respondents be conducted periodically to assess non-response bias in a particular injured worker population.

What about Spanish-speaking workers?

The survey will be available in Spanish by summer. The Spanish version should be made available on request to the administering organization, as indicated in the cover letter.

Future DWC activities

DWC has received funding from the Robert Wood Johnson to conduct a telephone survey using the revised patient satisfaction survey instrument. The survey is currently

being administered to nearly 800 injured workers from four populations, including the 24-hour pilot project, certified HCOs, large self-insured employers, and non-HCO workers' compensation managed care programs.

Many organizations have already expressed interest in using the DWC survey. We are exploring the possibility of creating a mechanism for pooling information from various survey users to create benchmark data for patient satisfaction in California's workers' compensation system. DWC will host a meeting to discuss the use of the survey later this spring. If you are interested in participating, please call Ms. Kathy Dervin, at (415) 703-4651.

References

1. California Labor Code. 4600.5.
2. California Code of Regulations. Title 8. Certification Standards for Health Care Organizations. . 1996.
3. AHCPR. Consumer Assessment of Health Plans.
4. Miller R LH. Managed Care Performance: Is Quality of Care Better or Worse? *Health Affairs*. 1997;16(5).
5. Allen HM Jr RW. Consumer surveys of health plan performance: a comparison of content and approach and a look to the future. *Joint Commission Journal on Quality Improvement*. 1996;22:775-94.
6. NCQA. HEDIS 3.0.(Vol.3) Member Satisfaction Survey: Protocols. Jan.1997, Wash.D.C.
7. Ware J SC. The MOS 36-Item Short-Form Health Survey (SF-36)" 1. Conceptual framework and item selection. *Medical Care*. 1992;30:473-481.
8. Cheadle A. Franklin G, Wolfhagen C, et.al. Factors influencing the duration of work-related disability. Am J Public Health 1994; 84:190-6. .
9. Franklin GM, Fulton-Kehoe D. Outcomes Research in Washington State Workers' Compensation. Am J Ind Med 1996; 29:642-8 .
10. Pransky G, Feuerstein M, Himmelstein J, et.al. Measuring functional outcomes in work-related upper extremity disorders. J Occup Environ Med. 1997; 39: 1195-1202.
11. Department of Labor and Industries University of Washington, Department of Health Services. Workers' Compensation Managed care Pilot Project: Final Report to the Legislature. Olympia, Wa; 1997.
12. Kyes K FG, Weaver M. Reliability and Validity of Medical Outcome and Patient Satisfaction Measures Among Injured Workers in Washington State: A Pretest. *American Journal of Industrial Medicine*. 1997;31:427-434.
13. Borba P AD, Fung M. Workers' compensation managed care pilot project: Final report to the Florida legislature by the Florida Department of Insurance. Talahassee, Fl.: Milliman and Robertson; 1994.
14. Levy B, Wegman D, ed. Occupational Health. 1983. Little, Brown and Co. Boston. .
15. Sum J. Navigating the Workers' Compensation System: The Injured Worker's Experience. Labor Occupational Health Program, UC Berkeley and California Commission on Health and Safety and Workers' Compensation. 1997.
16. Goyder J. *The Silent Minority: Non-respondents in Sample Surveys*: Polity Press; 1987.
17. Sherbourne CD HR, Burton T. Population-based surveys of access and consumer satisfaction with health care. *Consumer Survey Information in a Reforming Health Care System.* Bethesda: Agency for Health Care Policy and Research; 1995.
18. Walker AH RJ. Obtaining information of patient satisfaction with hospital care: mail vs. telephone. *Health Services Research*. 1984;19:291-296.

19. Mizes JS FE, Roos C. Incentives for increasing return rates: magnitude levels, response bias, and format. *Public Opinion Quarterly*. 1984;48:794-800.